RNA profiling in host–pathogen interactions

Simon J Waddell¹, Philip D Butcher¹, and Neil G Stoker²
¹Medical Microbiology, Division of Cellular & Molecular Medicine, St George’s University of London, Cranmer Terrace, Tooting, London, SW17 0RE, UK
²Pathology and Infectious Diseases, Royal Veterinary College, Royal College Street, London NW1 0TU, UK

Abstract
The development of novel anti-bacterial treatment strategies will be aided by an increased understanding of the interactions that take place between bacteria and host cells during infection. Global expression profiling using microarray technologies can help to describe and define the mechanisms required by bacterial pathogens to cause disease, and the host responses required to defeat bacterial infection.

Introduction
Exploring the RNA profiles of both host and pathogen through the course of infection promises to illuminate much about the infectious process and aid in the development of successful treatment strategies. This review focuses on the advances in whole genome transcriptional profiling of bacterial pathogens and host cells within the contexts of tissue culture, animal model and human disease. The bacterial transcriptional response to infection offers insight into the physiological state of infecting bacteria and the mechanisms required by bacteria to successfully survive infection [1]; this information could be used to define novel drug development strategies. The intracellular bacterial transcriptome might also be exploited as a bioprobe, to describe the microenvironments encountered by bacteria through the course of infection [2,3]. Additionally RNA profiling might be employed to identify novel vaccine candidates [4]. The transcriptional response of host cells to bacterial infection enables the intracellular and intercellular interactions to be explored throughout disease progression, facilitating the discovery of bacterial immunomodulatory actions. Furthermore the transcriptional signature of human non-invasive samples also promises to reveal novel diagnostic or predictive applications [5]. Here we review the recent advances in whole genome transcriptional profiling of both host and pathogen in ex-vivo, animal model and human disease contexts. Earlier reviews have been published in this field and might be found useful [6,7,8].

RNA methodologies
The ability to successfully monitor changes in transcript abundance is dependent on the development of RNA extraction techniques capable of purifying representative RNA populations from a variety of disease settings. The use of RNA stabilizing solutions is paramount in prokaryotic expression analyses, as the bacterial transcriptional response to the extraction process might mask relevant changes in gene expression [9]. An additional problem in studying the bacterial transcriptome during infection is the requirement to
separate eukaryotic from bacterial gene expression patterns; this is especially important in paucibacillary infections where the specific activity of labelled bacterial cDNA will be low relative to the background of host cDNA in a total RNA extraction reaction. Four strategies have been employed to overcome this problem, all of which, if validated correctly, enable the transcriptional response of bacteria from mixed RNA populations (host and pathogen) to be examined.

First is the development of differential lysis methods of bacterial RNA extraction, whereupon the host cells are lysed, whilst the bacterial cells remain intact to be recovered for RNA extraction, thereby enriching the bacterial mRNA several thousand-fold. This method was originally developed for Mycobacterium tuberculosis [3**,10,11] and a modified differential lysis approach has similarly been used for Salmonella [12*]. For a review about RNA extraction issues during infection, see [13]. Second is the use of negative selection methods to remove eukaryotic RNA from a population of mixed total RNA [14]. The third involves selective capture hybridisation (SCOTS) strategies which select for specific prokaryotic message [15] and the fourth is the utilization of DNA microarrays to discriminate between bacterial and host transcriptional profiles [16**].

The study of eukaryotic and bacterial mRNA populations has been enabled for individual genes by techniques such as northern blotting, SAGE, nuclease protection, primer extension, in situ hybridisation and particularly quantitative reverse-transcription PCR (RTq–PCR using chemistries such as Taqman or molecular beacons). More population-based mRNA analysis was facilitated before genome sequence availability by methods such as RNA arbitrarily primed (RAP)-PCR and differential display (DD)-RT–PCR; now, however, whole genome expression levels can be simultaneously measured using microarray technology. This review details the impact of whole genome expression profiling on the study of host–pathogen dynamics during infection.

Complexities of host–pathogen models

Clearly the transcriptional response to infection measured is dependent on the system investigated; the interactions between infecting bacteria and the complex mixture of cell types in vivo are likely to be different to those with a single cell-type cultured in vitro. The simplistic interface of host and pathogen in vitro might not reflect the heterogeneous cell types and microenvironments encountered in vivo, but can be used to define bacterial interactions with key cell types that provide a mechanistic insight into bacterial disease progression. Moreover, interpreting the gene expression data from mixed tissues or from bacteria in multiple micro-environments, as would be seen in complex tissues with bacteria in each location expressing different sets of overlapping genes, poses a considerable hurdle to understanding the complex network of interactions that occur during human disease. These datasets might however provide a global perspective of infection, highlighting diagnostic or predictive gene expression signatures.

Here, we divide the host–pathogen transcriptional literature into three sections dependent on the infection model used, from in vitro tissue culture studies using cell lines or extracted primary cells, to animal models of infection, and to human patient samples (Figure 1).

Tissue culture or primary extracts

Advantages

Many of the global expression analyses have been performed on cell lines or primary extracts in ex vivo model systems. This is largely a result of three factors: the ability to extract sufficient bacterial RNA from in vitro infection models; the availability of host cell types or primary extracts for multiple timepoints/comparisons; and the opportunity to
explore the relatively ‘simple’ interaction between a single bacterial species and a fixed cell type or cellular composition.

**Inferring bacterial metabolic states and defining virulence mechanisms**

The transcriptional patterns of intracellular bacteria have been used to define the responses required for survival and successful infection. Schnappinger et al. [3**], using a differential lysis method for RNA extraction, described the switch in *M. tuberculosis* metabolism from using glucose and glycerol as a carbon source to using fatty acids in the murine macrophage, together with an induction of genes implicated in the adoption of an anaerobic respiratory state. Comparison with in vitro transcriptomes under defined conditions reveals that the bacterial intracellular transcriptome can act as a bioprobe for the intracellular compartment in which the pathogen resides, which for *M. tuberculosis* suggested that the endosome is low in iron, oxidative, nitrosative and functionally hypoxic [3**]. The intracellular transcriptomes of *Salmonella enterica* [12•] and *Shigella flexneri* [17•] have also been described after phagocytosis of macrophage-like cell lines using differential lysis methodologies, and indicate that genes associated with type III secretion systems appear to be repressed on infection. The bacterial reaction to the intracellular environment also suggests that both magnesium and phosphate are limited during macrophage infection, but interestingly for *Salmonella* the vacuole was not limiting for iron [12•]. The significance of bacterial type III secretion systems has also been investigated by defining the response of murine macrophage or epithelial cell lines to infection with *Yersinia enterocolitica* [18] or *Pseudomonas aeruginosa* [19]. Additionally the role of *Helicobacter pylori* type IV secretion systems at the interface with gastric epithelial cells has been demonstrated [20]. Belland et al. [14] defined the transcriptional pattern associated with *Chlamydia trachomatis* growth in epithelial cells by removing the growth-inhibitory effect of interferon-γ (IFN-γ) from the culture medium. The relationship between global mediators of macrophage activation and *M. tuberculosis* infection has been investigated by Shi et al. [21]. The shift to an alternate metabolic state has also been characterised after macrophage or epithelial cell infection by *Listeria monocytogenes* [22,23]. Conversely, McCaffrey et al. [24] have identified a cluster of interferon-responsive genes induced in murine macrophages after infection with cytosol-localized compared to vacuole-restricted *L. monocytogenes*.

**Cell adhesion and invasion**

The changing pattern of bacterial gene expression might also be used to identify factors required for cell adhesion and entry; Dietrich et al. [25] analysed the transcriptome of *Neisseria meningitidis* during epithelial and endothelial infection, identifying genes that were differentially regulated in a single cell type only. Similarly Jain et al. [26] demonstrated the induction of a cluster of *M. tuberculosis* genes of unknown function involved in the invasion of brain endothelial cells, necessary if *M. tuberculosis* is to cross the blood–brain barrier in central nervous system (CNS) infection. Factors affecting successful cell entry might also be characterised by following host cell responses; Pedron et al. [27•] compared the expression pattern of epithelial cells after infection with invasive or non-invasive *S. flexneri* strains.

**Pathogen clearance or survival**

Microarray analysis has also been used to investigate why some bacterial pathogens are not eradicated successfully from blood; Voyich et al. [28] identified genes involved in capsule biosynthesis and oxidative stress as induced in *Staphylococcus aureus* in response to human polymorphonuclear leukocytes. In addition, several genes of unknown function were observed to be differentially regulated in strains more resistant to killing.
Host responses

Whole genome transcriptomics of the host has enabled pathogen-specific gene expression responses to be recognized in purified or complex cellular environments [29–31]. Comparative microarray analyses enable distinct transcriptional responses to be characterised, dependent on the infecting bacterium [32], highlighting potential immunomodulatory features such as the limited interleukin-12 (IL-12) production in macrophages infected with *M. tuberculosis* [33]. This comparative approach also reveals differences in the way host immune cells respond to the same infecting bacteria [34]; Granucci *et al.* [35] demonstrated that IL-2 is produced by dendritic cells but not macrophages after *Escherichia coli* infection. The transcriptional profiles of host cells in response to infection has been reviewed recently by Jenner and Young [6*].

Animal models

Animal models of infection can be used to profile host and pathogen transcriptomes in complex environments which cannot be recreated *in vitro* and for which human samples are largely unavailable. Interpretation of the RNA profiles is dependent primarily on the relevance of the animal system chosen.

Complexities of multi-cellular tissue environments

The rabbit ileal loop model of infection has been used to explore the transcriptional responses of *Vibrio cholerae* [36] and *Campylobacter jejuni* [37] to the intestinal environment. Both studies characterize the environment encountered to be nutrient-limiting and oxygen-limiting and identify putative virulence genes induced by bacteria in the rabbit intestine. The expression profiling of host tissues after bacterial infection is complicated by the changing cellular composition of organs after bacterial infection, and the selection of suitable control samples. Huff *et al.* [38], however, used biopsies from the antrum and corpus of nonhuman primates to describe the gastric transcription pattern through the course of *H. pylori* infection. The application of laser microdissection microscopy permits distinct cellular populations to be separated from complex tissues; Mueller *et al.* [39] compared the responses of parietal, mucus-producing and chief cell epithelial lineages in the murine stomach to *H. pylori* infection and demonstrated that a response to *H. pylori* is only detected in the mucus-producing cell type.

The environmental niche that bacteria occupy through the course of disease often defines the animal model selected for expression profiling. The gene expression pattern of the syphilis spirochete *Treponema pallidum*, which is intractable to RNA profiling *in vitro*, has been described in rabbit testicular tissue [40], and the importance in *Y. pestis* of a protective response to reactive nitrogen species was demonstrated in the rat bubo, the disruption of which causes attenuated virulence [41]. Indeed, *Y. pestis* virulence genes involved in type III secretion have been identified to be induced in the murine lung [42].

The response of *M. tuberculosis* has also been profiled in murine lung tissue, and Talaat *et al.* [43] also compared the expression pattern of *M. tuberculosis* in the lungs of immune-competent and immune-deficient mice. This approach has also been adopted to investigate lipoprotein expression by *Borrelia burgdorferi* in mice [44].

Advantages of animal models

The use of animal models enables three aspects of infection to be investigated that cannot easily be assessed by alternative methods: the interaction of bacteria with complex environments; the comparison of bacterial and/or host expression profiles from different sites of infection, for example Orihuela *et al.* [45] looked at the differential regulation of
Streptococcus pneumoniae genes in whole blood and cerebrospinal fluid (CSF), and identified tissue-specific expression of bacterial genes; and the correlation of bacterial or host RNA profiles with clinical parameters through the course of infection [46].

Problems with low bacterial abundance

The low number of bacteria and the difficulty in isolating bacterial RNA from host tissue has lead to the development of several systems designed to contain bacteria in an in vivo environment. Yarwood et al. [47] used subcutaneously implanted perforated hollow golf balls to model the S. aureus adaptive response, Karakousis et al. [48] used a hollow fiber assay to describe the interaction between M. tuberculosis and the artificial murine granuloma created, and the study of B. burgdorferi transcriptomics during mammalian infection has been aided by the use of implanted dialysis membrane chambers [49].

Human disease

Reports of global gene expression profiling in human tissue or non-invasive patient samples suffering from bacterial disease are understandably limited. Rachman et al. [50] have described the M. tuberculosis transcriptome in lung tissues extracted during surgery for untreatable tuberculosis, and were able to compare gene expression signatures of M. tuberculosis in the granuloma, pericavitial tissue and macroscopically normal lung. The gene expression profile of V. cholerae in human stool samples reinforces the model that this organism reaches a hyper-infectious state after colonisation of the human intestine, with genes involved in nutrient acquisition and motility induced [51]; indeed V. cholerae transcriptomes from human stool and vomitus have been compared to model the changing expression pattern from early to late stage infection [52]. The transcriptional profiles of human gastric biopsies before and after elimination of H. pylori infection were compared after laser capture microdissection by Resnick et al. [53], revealing the differential expression of established virulence genes and genes of unknown functional significance in H. pylori disease. The global expression profiling of host responses in patient blood [54] or saliva [55] has great diagnostic potential, as well as affording the opportunity to define host gene expression patterns in primary settings. Furthermore Ramilo et al. [56] have recently identified transcriptional signatures able to discriminate between E. coli and S. aureus infection in patient peripheral blood leukocytes.

Conclusions

Much of the power of transcriptome studies to identify changes in global gene expression patterns comes from the ability to compare RNA profiles from different bacterial strains and/or cell types or microenvironments. To this end, transcriptional data from in vitro studies detailing the responses of bacteria/host cells to different conditions/treatments enables the complex in vivo patterns of gene expression to be interpreted. A microarray experimental strategy, including the use of a reference control channel (such as genomic DNA in bacterial systems) and the adoption of ‘minimum information about a microarray experiment’ (MIAME)-compliant microarray databases [57] has also helped in this respect. The comparative nature of microarray analyses however also raises the question of what is the appropriate control RNA to compare in vivo transcriptional data against? Should bacterial in vivo expression data be compared to RNA extracted from in vitro logarithmic grown bacteria, bacteria resuspended in culture medium, or washed off the infected cells? How are differences in the cellular composition of infected compared to uninfected samples accounted for in eukaryotic transcriptional data, and do they need to be? A variety of strategies have been employed, dependent on the aspect of the host–pathogen relationship to be explored. The comparative biasing of microarray data, however, must be considered as more transcriptional datasets are generated and compared.
The development of bacterial amplification techniques [16**] and laser microdissection microscopy promises to ensure that the transcriptional profiling of distinct and previously intractable host–pathogen interactions continues to play an important role in understanding bacterial disease processes.

Acknowledgments

SJW was funded by an EU Sixth Framework Programme Priority (LHP-CT-2004-012187). We would like to acknowledge the Wellcome Trust and its Functional Genomics Resources Initiative for funding the multi-collaborative microbial pathogen microarray facility at St George’s (B@G@S).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

• of special interest

•• of outstanding interest


27. Pedron T, Thibault C, Sansonetti PJ. The invasive phenotype of Shigella flexneri directs a distinct gene expression pattern in the human intestinal epithelial cell line Caco-2. J Biol Chem. 2003;
278:33878–33886. [PubMed: 12813033] Comparison of the host response to invasive and non-invasive S. flexneri strains was used to define invasion-responsive genes in epithelial cells.


Figure 1.
An illustration of the options available to study the RNA profiling of host–pathogen interactions and the alternative perspectives offered.